

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: August 18, 2002, 08:18:46 : Search time 61.25 Seconds  
(without alignments)  
865.016 Million cell updates/sec

Title: US-09-777-921A-2  
Perfect score: 2491  
Sequence: 1 MLRWLRDFALPTAACQDAEQ.....VGISVYVENMKOTLGVTKQ 477

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2471	99.6	477	22	AAW79077 Human protein SEQ
2	1552.5	62.6	471	22	AAU27697 Human full-length
3	1552.5	62.6	509	22	AAU27869 Human contig polyp
4	1501	60.5	469	22	AAU66718 Membrane-bound pro
5	1501	60.5	469	22	AAU67554 Human PRO1106. HO
6	1501	60.5	469	22	AAU65241 Human PRO1106 (UNQ
7	1499	60.4	469	22	AAU76084 Murine ADP/ATP tra
8	1499	60.4	469	22	AAU56023 Skin cell protein,
9	1479.5	59.6	508	22	ABG22637 Novel human diagno
10	1394	56.2	385	21	AAU42129 Human ORFX ORF2093
11	1325	53.4	365	22	AAU40072 Human polypeptide

12	1241	50.0	461	22	AG65800 Human membrane tra
13	1143	46.1	244	22	AAU43524 Human polypeptide
14	1143	46.1	244	22	AAU19941 Novel human calciu
15	1037	41.8	292	21	AAU50388 Human uncoupling p
16	910	36.7	312	21	AAU75964 Murine skin cell p
17	910	36.7	312	22	AAU55903 Skin cell protein,
18	902	36.4	370	22	ABU71306 Drosophila melanog
19	841	33.9	182	22	AAU80051 Human protein SEQ
20	800.5	32.3	226	22	AAU41858 Human polypeptide
21	761	30.7	208	22	AAU93247 Human polypeptide
22	632	25.5	169	21	AAU50389 Human uncoupling p
23	521	21.0	352	21	AAU30070 Arabidopsis thalia
24	508	20.5	241	21	AAU12158 Arabidopsis thalia
25	508	20.5	330	21	AAU28434 Arabidopsis thalia
26	508	20.5	332	21	AAU28433 Arabidopsis thalia
27	505.5	20.4	166	21	AAU42319 Human ORFX ORF2083
28	479	19.3	138	22	ABG22634 Novel human diagno
29	455.5	18.4	415	21	AAU29005 Arabidopsis thalia
30	454.5	18.3	289	20	AAU31936 Wheat brittle-1 pa
31	454.5	18.3	381	21	AAU29257 Arabidopsis thalia
32	454.5	18.3	384	21	AAU29256 Arabidopsis thalia
33	446	18.0	392	21	AAU11515 Arabidopsis thalia
34	446	18.0	392	21	AAU49411 Arabidopsis thalia
35	446	18.0	411	21	AAU11514 Arabidopsis thalia
36	446	18.0	411	21	AAU49410 Arabidopsis thalia
37	445	17.9	291	21	AAU30071 Arabidopsis thalia
38	432.5	17.4	316	21	AAU11516 Arabidopsis thalia
39	432.5	17.4	316	21	AAU49412 Arabidopsis thalia
40	417	16.8	198	21	AAU12159 Arabidopsis thalia
41	412	16.6	266	21	AAU28435 Arabidopsis thalia
42	410.5	16.5	377	22	ABU60506 Drosophila melanog
43	382.5	15.4	273	21	AAU29006 Arabidopsis thalia
44	380.5	15.3	267	21	AAU29258 Arabidopsis thalia
45	369	14.9	323	22	AAU41127 Human polypeptide

## ALIGNMENTS

## RESULT 1

AAW79077  
ID AAW79077 standard; Protein: 477 AA.

XX AC AAW79077;  
XX

DT 06-NOV-2001 (first entry)  
XX

DE Human protein SEQ ID NO 1739.  
XX

DE Human: cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.

XX OS Homo sapiens.  
XX

XX PN WO200157190-A2.  
XX

XX PD 09-AUG-2001.  
XX

XX PF 05-FEB-2001; 2001WO-US04098.  
XX

XX PR 03-FEB-2000; 2000US-0496914.  
XX

XX PR 27-APR-2000; 2000US-0560875.  
XX

XX PR 20-JUN-2000; 2000US-0598075.  
XX

XX PR 19-JUL-2000; 2000US-0620325.  
XX

XX PR 01-SEP-2000; 2000US-0654936.  
XX

XX PR 15-SEP-2000; 2000US-0663561.  
XX

XX PR 20-OCT-2000; 2000US-0693325.  
XX

XX PR 30-NOV-2000; 2000US-0728422.  
XX

XX PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
 XX WPI: 2001-476283/51.  
 DR N-PSDB; AAK52210.  
 XX Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX Claim 20; Page 4074-4075; 6221pp: English.  
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAK78333-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAK80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX Sequence 477 AA;

Query Match 99.6%; Score 2471; DB 22; Length 477;  
 Best Local Similarity 99.6%; Pred. No. 1.8e-217;  
 Matches 475; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 MLRLWDLFALPTAACODAEOPRYETLFOALDRNGDGVWDIGELQELNLTGIPGODAE 60  
 DB 1 mlrwldfpltaacodaeopryetlfgaldngdgvwdigelnltgipgoda 60  
 OY 61 EKIFTTGDNVKGKLDPEEFMKYLDHEKMKLAFKSLDKNNNGKTEASEIVOSLTGL 120  
 DB 61 ekifftgdnvkgldpeefmkylkdhkmmkklafksldknnngkteseivogslgtlgl 120  
 OY 121 TISEQQAELIQLSDVDGTFVQNMENRDYFLFNPVTDISEIRFMKHSKGIDIGDSLTI 180  
 DB 121 tiseqqaellqlsldvdgtfvtvqnmernrdyflfnpvtidiseirfmkhskgidigdslti 180  
 OY 181 PDEFTDEKKSQWROLLAGGTAGAVSRTSTAPLRLKTMQVHCSSDKMNITGGERQ 240  
 DB 181 pdeftekekksqwrllaggtagavsrststaprlrlktmqvhscksdmknitggrq 240  
 OY 241 MVKGGIRSLWRNGTNNVITAPETAVKFWAYBOYKLLLTBEGQKGTGTFERFISGMAGA 300  
 DB 241 mvkggirslwrgntnnvitiapetavkfwayboykllltbeggkgtgtgterfismaga 300  
 OY 301 TAQTFIYPMVEMKTRVLAVGKTQVSGIYDCAKKILKHEGIGAFYKGVNPLGIIPYAGI 360  
 DB 301 taqtfiypmvmktrvlavgktqvsgiydcakklkhegigafykgvnpnlgiiipyagi 360  
 OY 361 DLAVYELLKSWLONFAKDSVNPQVWVLLCGALSTSCQLASYPALVTRMQAAMLE 420  
 DB 361 dlavyellkswlonfakdsvnqvwvllcgalsstscqlasypalvtrmqaaamle 420  
 OY 421 GSPQLNMVGLFRRRIISKEGIPGLYRGITPDMFKVLPAVGISYVYVNNKQTLGVTOK 477  
 DB 421 gspqlnmvglfrrriiskegipglyrgitpdmfkmvlpavgisyyvynnkqtlgtvok 477

RESULT 2

AAU27697

AAU27697 standard; Protein: 471 AA.

AAU27697;

18-DEC-2001 (first entry)

Human full-length polypeptide sequence #22.

Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
 mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
 cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
 nervous system disorder; inflammatory disorder; cell differentiation;  
 angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;  
 genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
 cytostatic; antirheumatic; antiarthritic; vulnery; antinflammatory;  
 antibacterial; immunosuppressive; vasotropic; antiparkinsonia;  
 neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;  
 immunostimulant; analgesic; gene therapy.

Homo sapiens.

WO200164834-A2.

07-SEP-2001.

26-FEB-2001; 2001WO-US04926.

28-FEB-2000; 2000US-0515126.

18-MAY-2000; 2000US-0577409.

17-JUN-2000; 2000US-0597707.

14-JUL-2000; 2000US-0616807.

19-SEP-2000; 2000US-0664641.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 Drmanac R;

WPI: 2001-589862/66.

N-PSDB; AAK544597.

Novel polypeptides and nucleic acids obtained from cDNA libraries  
 prepared from various human tissues, for diagnosis, treatment of  
 cancer, neurological, inflammatory disorders and for use in arrays for  
 detection -

Claim 10; SEQ ID NO 194; 153pp; English.

Sequences AAU27676-AAU28019 represent full-length polypeptides and  
 contig polypeptides of the invention. The proteins and their associated  
 DNA sequences are useful for the treatment, diagnosis and prevention of  
 various types of disorder in a mammalian subject such as a human, dog,  
 monkey, mouse, hamster or rat. The disorders include cancers such as  
 leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as  
 multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 bowel disease. The sequences exhibit activity relating to angiogenesis,  
 cell proliferation, cell differentiation, stem cell growth factor,  
 activin or inhibin. Therefore, they can be used to manipulate stem cells  
 in culture to give rise to neuroepithelial cells that can be used to  
 augment or replace cells damaged by illness, accidental damage or genetic  
 disorders. The sequences may also be used for regeneration of bone,  
 cartilage, tendons and ligaments and in tissue repair and burn healing.  
 Note: Some sequences for this patent did not form part of the printed  
 specification, but were obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_sequences.

Sequence 471 AA;

Query Match 62.6%; Score 1552.5; DB 22; Length 471;  
 Best Local Similarity 61.5%; Pred. No. 2.3e-133;

Matches 286; Conservative 85; Mismatches 87; Indels 7; Gaps 2;

Oy 17 DAEOPTRYETLFOALDRNGGVVDIGELQBLRNGLIPGLQD----AEEKIFTGDNVKD 72  
 Db 7 daerrqrwgrlfealdankdgrvvhelrglarlg---ggnpdpgagqgissegdadp 63  
 Oy 73 KGLDFEPMYKLDHEKMKLAFKSLDKNDGKIEASEIVQSLOTGLTISQOAEILQ 132  
 Db 64 ggldeefsyrlqereqrlllmfhsldrnqghidvseiqsfalgisileqaekilh 123  
 Oy 133 SIDVGTWTVDWNEWRDYFENPVTDEIIRFWKHSTGIDIGDSTLTPDEFTDEKSG 192  
 Db 124 smdrdgtmdidwqewdhflhslenvedvlyfwhstvdldgecltvpdefskqektg 183  
 Oy 193 QMWRQLLAGGIAGAVSRTSTAPDLRLKIMNOVHGSKSDKMNIFGFRWVVEGGIRSLWR 252  
 Db 184 mwwqlvagavagavartgtcapldrlkvfmqvhasktnrlnlgslrsmvleggirsrlwr 243  
 Oy 253 GNGTNVTKIAPETAVKFWAYEQYKLLTBEGQKIGTFERFISGSMAGATATQTFYPMVEM 312  
 Db 244 gnginviklapesaifmayeqikrallgqgetlhvqerfvagslagataqtilypmevl 303  
 Oy 313 KTRLAVGKTQYSGIYDCAKILKHEGLGAFYKGVVNPVNLGIPVAGIDLAVYELLKSYW 372  
 Db 304 ktrltirttqykgllidcarrileregprafyrgyilpvnvlgilpvgidlavetlknw 363  
 Oy 373 LDNFAKDSVNPVWVLLGCGLSSTCGQLASYPALVTRRMOAQLGSPQLNMVGLFR 432  
 Db 364 lqyshdsadpgilvillacgtlastcgqlasyplalvtrmqagasteggpqslmlgllr 423  
 Oy 433 RIISKEGIPGLYRGITPNNKVLPAVGISYVYVYNNKOTLGVTK 477  
 Db 424 hllsqgmrglyrglapnmkvpavsisvvyenmkqelgytar 468

RESULT 3

AAU27869 standard; Protein: 509 AA.

AAU27869;

18-DEC-2001 (first entry)

Human contig polypeptide sequence #22.

Mammal: human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
 mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
 Cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
 nervous system disorder; inflammatory disorder; cell differentiation;  
 angiogenesis; stem cell growth factor; actinin; inhibin; cartilage; burn;  
 genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
 cytostatic; antitumor; antirheumatic; antiarthritic; vulnary; antinflammatory;  
 antibacterial; immunosuppressive; vasotropic; antiparkinsonian;  
 neuroprotective; osteopathic; antidiabetic; antiallergic;  
 immunostimulant; analgesic; gene therapy.

OS Homo sapiens.

OS Synthetic.

XX WO200164834-A2.

PN 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04936.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PR 17-JUN-2000; 2000US-0507707.

PR 14-JUL-2000; 2000US-0616807.

PR 19-SEP-2000; 2000US-0664641.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 PI Drmanac R;

DR WPI: 2001-589862/66.  
 DR N-PSDB: AAS44769.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues for diagnosis, treatment of  
 PT cancer, neurological, inflammatory disorders and for use in arrays for  
 PT detection

PS Claim 10: Page 126-127; 153pp; English.

XX Sequences AAU27676-AAU28019 represent full-length polypeptides and  
 CC contig polypeptides of the invention. The proteins and their associated  
 CC DNA sequences are useful for the treatment, diagnosis and prevention of  
 CC various types of disorder in a mammalian subject such as a human, dog,  
 CC monkey, mouse, hamster or rat. The disorders include cancers such as  
 CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as  
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 CC bowel proliferation, cell differentiation, stem cell growth factor,  
 CC cell proliferation, cell differentiation, stem cell growth factor,  
 CC actinin or inhibin. Therefore, they can be used to manipulate stem cells  
 CC in culture to give rise to neuroepithelial cells that can be used to  
 CC augment or replace cells damaged by illness, accidental damage or genetic  
 CC disorders. The sequences may also be used for regeneration of bone,  
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.  
 CC Note: Some sequences for this patent did not form part of the printed  
 CC specification, but were obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 509 AA;

Query Match 62.6%; Score 1552.5; DB 22; Length 509;

Best Local Similarity 61.5%; Pred. NO. 2.6e-133;

Matches 286; Conservative 85; Mismatches 87; Indels 7; Gaps 2;

Oy 17 DAEOPTRYETLFOALDRNGGVVDIGELQBLRNGLIPGLQD----AEEKIFTGDNVKD 72  
 Db 45 daerrqrwgrlfealdankdgrvvhelrglarlg---ggnpdpgagqgissegdadp 101  
 Oy 73 KGLDFEPMYKLDHEKMKLAFKSLDKNDGKIEASEIVQSLOTGLTISQOAEILQ 132  
 Db 102 ggldeefsyrlqereqrlllmfhsldrnqghidvseiqsfalgisileqaekilh 161  
 Oy 133 SIDVGTWTVDWNEWRDYFENPVTDEIIRFWKHSTGIDIGDSTLTPDEFTDEKSG 192  
 Db 162 smdrdgtmdidwqewdhflhslenvedvlyfwhstvdldgecltvpdefskqektg 221  
 Oy 193 QMWRQLLAGGIAGAVSRTSTAPDLRLKIMNOVHGSKSDKMNIFGFRWVVEGGIRSLWR 252  
 Db 222 mwwqlvagavagavartgtcapldrlkvfmqvhasktnrlnlgslrsmvleggirsrlwr 281  
 Oy 253 GNGTNVTKIAPETAVKFWAYEQYKLLTBEGQKIGTFERFISGSMAGATATQTFYPMVEM 312  
 Db 282 gnginviklapesaifmayeqikrallgqgetlhvqerfvagslagataqtilypmevl 341  
 Oy 313 KTRLAVGKTQYSGIYDCAKILKHEGLGAFYKGVVNPVNLGIPVAGIDLAVYELLKSYW 372  
 Db 342 ktrltirttqykgllidcarrileregprafyrgyilpvnvlgilpvgidlavetlknw 401  
 Oy 373 LDNFAKDSVNPVWVLLGCGLSSTCGQLASYPALVTRRMOAQLGSPQLNMVGLFR 432  
 Db 402 lqyshdsadpgilvillacgtlastcgqlasyplalvtrmqagasteggpqslmlgllr 461  
 Oy 433 RIISKEGIPGLYRGITPNNKVLPAVGISYVYVYNNKOTLGVTK 477

Db 462 hilsqegmrglyrglepnfmkvpavslayvvvymkqalgvtar 506

RESULT 4

AAV66718  
ID AAY66718 standard; protein; 469 AA.

XX AAY66718;

XX 05-APR-2000 (first entry)

XX Membrane-bound protein PRO1106.

KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;  
pharmaceutical; receptor immunoadhesin; gene mapping.

XX Homo sapiens.

XX WO9963088-A2.

XX 09-DEC-1999.

XX 02-JUN-1999; 99WO-US12252.

XX 02-JUN-1998; 98US-0087607.

XX 02-JUN-1998; 98US-0087609.

XX 03-JUN-1998; 98US-0087759.

XX 04-JUN-1998; 98US-0088021.

XX 04-JUN-1998; 98US-0088025.

XX 04-JUN-1998; 98US-0088028.

XX 04-JUN-1998; 98US-0088029.

XX 04-JUN-1998; 98US-0088030.

XX 04-JUN-1998; 98US-0088033.

XX 05-JUN-1998; 98US-0088326.

XX 05-JUN-1998; 98US-0088167.

XX 05-JUN-1998; 98US-0088202.

XX 05-JUN-1998; 98US-0088212.

XX 05-JUN-1998; 98US-0088217.

XX 09-JUN-1998; 98US-0088655.

XX 10-JUN-1998; 98US-0088722.

XX 10-JUN-1998; 98US-0088730.

XX 10-JUN-1998; 98US-0088734.

XX 10-JUN-1998; 98US-0088738.

XX 10-JUN-1998; 98US-0088740.

XX 10-JUN-1998; 98US-0088742.

XX 10-JUN-1998; 98US-0088810.

XX 10-JUN-1998; 98US-0088811.

XX 10-JUN-1998; 98US-0088824.

XX 10-JUN-1998; 98US-0088825.

XX 11-JUN-1998; 98US-0088826.

XX 11-JUN-1998; 98US-0088858.

XX 11-JUN-1998; 98US-0088861.

PR 19-JUN-1998; 98US-0089952.  
PR 22-JUN-1998; 98US-0090246.  
PR 22-JUN-1998; 98US-0090252.  
PR 23-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 24-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 24-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092182.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095916.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096329.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 17-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096959.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.

PR 20-AUG-1998; 98US-0097218.  
 PR 24-AUG-1998; 98US-0097661.  
 PR 26-AUG-1998; 98US-0097951.  
 PR 26-AUG-1998; 98US-0097952.  
 PR 26-AUG-1998; 98US-0097953.  
 PR 26-AUG-1998; 98US-0097954.  
 PR 26-AUG-1998; 98US-0097955.  
 PR 26-AUG-1998; 98US-0097971.  
 PR 26-AUG-1998; 98US-0097974.  
 PR 26-AUG-1998; 98US-0097978.  
 PR 26-AUG-1998; 98US-0097979.  
 PR 26-AUG-1998; 98US-0097986.  
 PR 26-AUG-1998; 98US-0098014.  
 PR 31-AUG-1998; 98US-0098525.  
 PR 16-SEP-1998; 98US-0100634.  
 PR 12-JAN-1999; 99US-0115565.  
 XX (GETH ) GENE/TECH INC.  
 PA Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
 PI Wood WI, Yuan J;  
 PI WPI: 2000-072883/06.  
 DR N-PSDB; AA265058.  
 XX Membrane-bound proteins and related nucleotide sequences -  
 XX claim 12; Fig 206; 822pp; English.  
 XX The invention provides membrane-bound PRO polypeptides and  
 CC polynucleotides encoding them. The PRO sequences of the invention were  
 CC identified based on extracellular domain homology screening. The PRO  
 CC sequences have homology with proteins including LDL receptors, TIE  
 CC ligands and various enzymes. The membrane-bound proteins and receptor  
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
 CC immunoadhesins, for instance, can be used as therapeutic agents to block  
 CC receptor-ligand interactions. The membrane-bound proteins can also be  
 CC employed for screening of potential peptide or small molecule inhibitors  
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
 CC are useful as hybridization probes, in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
 CC will also be useful for the preparation of PRO polypeptides, especially  
 CC by recombinant techniques.  
 XX Sequence 469 AA;

Query Match 60.5%; Score 1501; DB 21; Length 469;  
 Best Local Similarity 65.8%; Pred. No. 1.2e-128;  
 Matches 275; Conservative 71; Mismatches 72; Indels 0; Gaps 0;  
 QY 60 EEKIFTGDNVNGKGLDFEEMKYLKDKHEKMKLAFKSLDKNDGKTEASEIVQSLOTGLG 119  
 DB 52 kqkivagdkldgldfeefvhyldhekkrlrvfildkndgridagelmsldlg 111  
 QY 120 LRISEQAELLLOSVDVGTVTVDNENROYFLENPNVTDEIRFWKXSTGIDGDSLT 179  
 DB 112 vxiseqaeelkismkngntidvnewrdyhlhlpvenipellykxhstifdvgenit 171  
 QY 180 IPDETEDEKSGOWMROLLAGGAGAVSRTSTAPDLKIMQVHGSKDKMNIFGGFR 239  
 DB 172 vpeftveerqtgmwrthlvagggagavstctcpldrklvlnqvhaarsnmglvggft 231  
 QY 240 QMVKEGGRLSWRGNGTVKIAPEATAVKFWAYEYOYKLLTETEGQIGTFERFISGMAG 299  
 DB 232 qmireggarslwrgnglnvklepsaikkmayeqkrlvgsdqetirherilvagslag 291  
 QY 300 ATAQTFTYPMEMKTRLVAGTGTGQYSGIYDCAKKILKHEGLGAFYKGYVPNLLGIIPVAG 359  
 DB 292 alagsiypmevltrmalrktgysgmldcarrilaregvaafykygypnmlgilpyag 351  
 QY 360 IDLAYELLKSYLDNFAKDSVNPVGVMLLCGALSSTCGOLASYPALVTRMQAQL 419  
 DB 352 idlayetlknawilqhyavnsadpgvfvllacgtmstcqqqlasyplavtrcmqagasi 411

QY 420 EGSPOLNMVGLFRRRIISKEGIPYRGITPNFMKVLPAVGISYVYVENKOTLGVTQK 477  
 DB 412 egapevtmssifkhlrtgafgiyrglapmfkvlpavslsyvvyenklitlgvqsr 469

## RESULT 5

AA887554  
 ID AAB87554 standard; Protein; 469 AA.

XX AC AAB87554;

XX DT 15-MAY-2001 (first entry)

XX DE Human PRO1106.

XX KW Human; PRO protein; mapping.

XX OS Homo sapiens.

XX PN WO200116318-A2.

XX PD 08-MAR-2001.

XX PF 24-AUG-2000; 2000WO-US23328.

XX PR 01-SEP-1999; 99WO-US20111.

XX PR 15-SEP-1999; 99WO-US21090.

XX PR 07-DEC-1999; 99US-0169495.

XX PR 09-DEC-1999; 99US-0170262.

XX PR 11-JAN-2000; 2000US-0175481.

XX PR 18-FEB-2000; 2000WO-US04341.

XX PR 22-FEB-2000; 2000WO-US04342.

XX PR 01-MAR-2000; 2000WO-US05601.

XX PR 03-MAR-2000; 2000US-0187202.

XX PR 25-APR-2000; 2000US-0199397.

XX PR 22-MAY-2000; 2000WO-US14042.

XX PR 05-JUN-2000; 2000US-0209832.

XX (GETH ) GENE/TECH INC.

XX Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;

XX WPI: 2001-183260/18.

XX N-PSDB; AAF92086.

XX Eighty four nucleic acids encoding PRO polypeptides, useful in  
 PT molecular biology, including use as hybridization probes, and in  
 XX chromosome and gene mapping.

XX Claim 12; Fig 58; 278pp; English.

XX The present sequence is a human PRO polypeptide (secreted and  
 CC transmembrane). The PRO protein, and PRO agonists, PRO antagonists or  
 CC anti-PRO antibodies are useful for preparation of a medicament useful in  
 CC the treatment of a condition which is responsive to the PRO protein,  
 CC agonists, antagonists or anti-PRO antibodies. The PRO protein may also be  
 CC employed as molecular weight markers for protein electrophoresis. The PRO  
 CC coding sequence has applications in molecular biology, including use as  
 CC hybridisation probes, and in chromosome and gene mapping.

XX Sequence 469 AA;

Query Match 60.5%; Score 1501; DB 22; Length 469;  
 Best Local Similarity 65.8%; Pred. No. 1.2e-128;  
 Matches 275; Conservative 71; Mismatches 72; Indels 0; Gaps 0;

QY 60 EEKIFTGDNVNGKGLDFEEMKYLKDKHEKMKLAFKSLDKNDGKTEASEIVQSLOTGLG 119  
 DB 52 kqkivagdkldgldfeefvhyldhekkrlrvfildkndgridagelmsldlg 111





Db 52 kqkivagdkldgldfeefvhyldhekkrlvfkldkndgridaqelmslrdlg 111  
 QY 120 LTISEQOAEILIQSDVGTWTVDMNWRDYFLPNVPTDIEETIRKWKSTGTGIDGST 179  
 Db 112 vkiseqqeeklikamdngtntidmewrdyhlhprvenpeillykwhstfidvgenit 171  
 QY 180 IDPETERDEKSKQWROLLAGGTAGAVSRSTAPLRLKIMQVHGSKSKDNIFGGR 239  
 Db 172 vpedtveerqgmwrhnlvaggagavsrctcpldrklvlnqvhasrnnmclvgft 231  
 QY 240 QMVKEGGRSLWRGNTVIRIAPETAVKFWAYEYOYFKLLTEGCKTGFPERFISGMAG 299  
 Db 232 qmireggakslvrgnglnvklapesaalkfmayeqmrlvgsdqetlrherlivagslag 291  
 QY 300 ATAOTFIYPMVEMKTRLAAGTGOYSGIYDCAKKILKHEGLGAFYKGYVNPULLGIIPYAG 359  
 Db 292 alaqslypmeviktrmairktgysgmldcarrilakegvaafykygypnmglipiyag 351  
 QY 360 IDLAVYELLSKWLONFANDSNVPMVLLGCGALSTCCQLASYPLALVTRMRQAQML 419  
 Db 352 idlavyetlknwlgryavnsadpgvfvllacgtlscstcgqlasyplalvtrmqasai 411  
 QY 420 EGSPOLNMVGLFRRIISKEGIPGLYRGITPNEMKVLPAVGISYVYENMKOTLGVTK 477  
 Db 412 egapevtnssifkqlrlrtegafglyrglapnfmkvpavsisyvvyenklitlgvqr 469

## RESULT 9

ABG22637  
 ID ABG22637 standard; Protein: 508 AA.

AC ABG22637;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #22628.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175057-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; RAS86824.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity

PS Claim 20; SEQ ID No 52996; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 508 AA;

Query Match 59.6%; Score 1479.5; DB 22; Length 508;  
 Best Local Similarity 59.3%; Pred. No. 1.3e-126;  
 Matches 277; Conservative 85; Mismatches 96; Indels 9; Gaps 4;

QY 17 DAEOPRYETLQALDRNGDGVWDIGELQELGNLGLPLGD-----AEKIFTGDNKD 72

Db 45 daerrgrwrlfeeldsnkdkgrvvhelrqglarlg---ggnpdpaqgqissegdapn 101

QY 73 GKLDPEEFMYLKDHEKKKLAFLAKSLDKNNDGKTEASEIVQSLQTLGLTISEQOAE-IL 131

Db 102 ggldleefarylqerqlmfmhaldnrgqghidvaeiqsraigisilleqaeknfa 161

QY 132 QSTDVDGTMVDNEMRDYELFNPVDTETETIRFNKHKSTCIDIGDSLTPDTEDEKRS 191

Db 162 qavrdgtmctidqewrdhflhseinvvedviyfwkhnstvidgecltvpdefskqekit 221

QY 192 GQWNROLLAGGIAGAVSRSTAPLRLKIMQVHGSKSKDNIFGGRQWVKEGGIRSLW 251

Db 222 gmwkqlaagavagavsrctcpldrklvfmqvhasktrlnllgslxsmvleggrrslw 281

QY 252 RGNQNVNVIKIPETAVKFWAYEYOYFKLLTEGCKTGFPERFISGMAGATAOTFIYPM 311

Db 282 rgnginvklapesaalkfmayeqkrlapqekitlpvlerfvagslagtaqtilypmav 341

QY 312 MKTRLAAGTGOYSGIYDCAK-KILKHEGLGAFYKGYVNPULLGIIPYAGIDLAVYELLS 370

Db 342 lktlritrttqykglaglrrggleregprafyrgypnvgilpyagidlavyetlkn 401

QY 371 YWLDNFAKSNVPMVLLGCGALSTCCQLASYPLALVTRMRQAQMLEGSPQLMAYGL 430

Db 402 wwlgqyshdsadpgllvllacgtlscstcgqlasyplalvtrmqasleggpklsmlgl 461

QY 431 FRRIISKEGIPGLYRGITPNEMKVLPAVGISYVYENMKOTLGVTK 477

Db 462 lrhlisqegmrglyrglapnfmkvpavsisyvvyenmkqalgvtar 508

## RESULT 10

AAB42329

ID AAB42329 standard; Protein: 385 AA.

AC AAB42329;

DT 08-FEB-2001 (first entry)

DE Human ORF2093 polypeptide sequence SEQ ID NO:4186.

KW Human; open reading frame; ORF; detection; cytostatic; hepatotropic;  
 KW vulnery; antipsoaric; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2002, 08:14:06 ; Search time 24.71 seconds  
(without alignments)  
471.510 Million cell updates/sec

Title: US-09-777-921a-2  
Perfect score: 2481  
Sequence: 1 MLRWLRDLPTRACDAEO.....VCISVYVYENKQTLGVYQK 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_MA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499	60.4	469	4	US-09-188-930-339
2	910	36.7	312	4	US-09-188-930-142
3	347	14.0	298	3	US-08-961-871-10
4	323.5	13.0	328	4	US-09-068-140A-15
5	311.5	12.6	289	4	US-09-068-140A-10
6	295	11.9	320	2	US-08-933-750C-12
7	295	11.9	320	4	US-09-234-613-12
8	281	11.3	674	4	US-09-160-119-2
9	269.5	10.9	447	4	US-09-160-119-4
10	255	10.3	351	2	US-08-933-750C-19
11	255	10.3	351	4	US-09-234-613-19
12	235.5	9.7	311	1	US-08-775-009-32
13	233.5	9.4	299	1	US-08-518-878B-56
14	233.5	9.4	309	1	US-08-470-868A-56
15	233.5	9.4	309	1	US-08-518-878B-51
16	233.5	9.4	309	2	US-08-470-868A-51
17	233.5	9.4	309	2	US-08-470-868A-51
18	233.5	9.4	309	3	US-09-210-681-51
19	233.5	9.4	309	3	US-08-946-719A-51
20	233.5	9.4	311	2	US-08-775-009-33
21	229	9.2	312	4	US-09-142-563-2
22	228	9.2	432	2	US-08-937-466-4
23	228	9.2	432	2	US-09-172-528-4
24	228	9.2	432	3	US-09-318-199-4
25	228	9.2	432	4	US-09-503-579-4
26	226	9.1	306	5	PCT-US94-09799-1
27	224	9.0	308	2	US-08-937-466-2

28	224	9.0	308	2	US-09-172-528-2	Sequence 2, Appl
29	224	9.0	308	3	US-09-318-199-2	Sequence 2, Appl
30	224	9.0	308	4	US-09-503-579-2	Sequence 2, Appl
31	198.5	8.0	303	1	US-08-518-878B-37	Sequence 37, Appl
32	198.5	8.0	303	2	US-08-294-522B-36	Sequence 36, Appl
33	198.5	8.0	303	2	US-08-807-861A-37	Sequence 37, Appl
34	198.5	8.0	303	2	US-08-470-868A-37	Sequence 37, Appl
35	198.5	8.0	303	3	US-09-210-681-37	Sequence 37, Appl
36	198.5	8.0	303	3	US-08-946-719A-37	Sequence 37, Appl
37	193.5	7.8	307	2	US-08-807-861A-56	Sequence 56, Appl
38	193.5	7.8	307	3	US-09-210-681-56	Sequence 56, Appl
39	193.5	7.8	307	3	US-08-946-719A-56	Sequence 56, Appl
40	186	7.5	125	4	US-08-905-223-320	Sequence 320, App
41	179	7.2	256	2	US-08-937-466-6	Sequence 6, Appl
42	179	7.2	256	2	US-09-172-528-6	Sequence 6, Appl
43	179	7.2	256	3	US-09-318-199-6	Sequence 6, Appl
44	179	7.2	256	4	US-09-503-579-6	Sequence 6, Appl
45	178	7.2	149	1	US-08-100-874-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-188-930-339  
: Sequence 339, Application US/09188930A  
: Patent No. 6150502  
: GENERAL INFORMATION:  
: APPLICANT: Watson, James D.  
: APPLICANT: Strachan, Lorna  
: APPLICANT: Sleeman, Matthew  
: APPLICANT: Chrust, Rene  
: APPLICANT: Murison, James Greg  
: TITLE OF INVENTION: Compositions Isolated From Skin Cells  
: FILE REFERENCE: 11000.1011c1  
: CURRENT APPLICATION NUMBER: US/09/188,930A  
: CURRENT FILING DATE: 1998-11-09  
: NUMBER OF SEQ ID NOS: 348  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 339  
: LENGTH: 469  
: TYPE: PRT  
: ORGANISM: Mouse  
US-09-188-930-339

Query Match	60.4%	Score 1499;	DB 4;	Length 469;
Best Local Similarity	65.6%	Pred No 5.5e-138;		
Mismatches	274;	Conservative	73;	Mismatches 71; Indels 0; Gaps 0;
QY	60	EKIFTGQVNDGKLDPEEFPMKYUKOHEKMKKLAFLSKLDKNDKDKIEASEIVOSLQTLG	119	
Db	52	KOKIVAGDKDGLDPEEFVYLDQHEKRLVFLSKLDKNDKNDGRIDAOEIMQSLRDLG	111	
QY	120	LTISEQAEILQSIDVQCTMTVDNNEWDYFLNPVPTDTEETIRFWKHSTGTDIGDSLT	179	
Db	112	VKISEQAEKILSKDKNGTNTIDNNEWDYHLLHPVENPEILYLYKHSITFDVGENLT	171	
QY	180	IPDETETEKKSGOWRLLAGGTAGAVSRSTAPDLKIMNOVHGSKSKDKNIFGGR	239	
Db	172	VDFEVEERQYGMWRHLVAGGAGAVSRCTAPDLRLKLVLMOVHASRNNHCIVGGFT	231	
QY	240	QMYEGGIRSLWRNGTNTVIKIAPETAVKFWAYDQYKLLTECOKIGTFRFISGSMAG	299	
Db	232	QMIREGGAKSLWRNGINVLKIAPEAIKFMAYEQMKRLVGSQDQTLRIHERLVAGSLAG	291	
QY	300	ATAQTFFYPMEVMTKTRLAGVTGOYSGIYDCAKKILKEGILGAFYKGYVNPILGIIPYAG	359	
Db	292	ATAQSSIIYPMEVLTNRALRTGQYSGMLDCARRILAKEGVAAYFYKGYIPNMLGIIPYAG	351	
QY	360	IDLAYTELLKSWLDONFAKOSVNPVGMVVLGCGGALSSTCCQLASYPLALVYTRMNOANML	419	



Db 419 CAPOLNMGVLFRRIRISREGLPLRGITPDMFKVLPVAGISVYVYENMKOTILGVTK 475

RESULT 2

T22688

hypothetical protein F55A11.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T22688

R:Kershaw, J.

submitted to the EMBL Data Library, May 1996

A:Reference number: 219600

A:Accession: T22688

A:Status: preliminary

A:Map position: 5

A:Gene: F55A11.4

A:Map position: 5

A:Introns: 24/1; 77/2; 277/3; 434/3; 474/1; 511/3; 531/3

Query Match 44.5% Score 1103; DB 2; Length 588;

Best Local Similarity 45.7% Pred. No. 3.9e-65;

Matches 210; Conservative 104; Mismatches 142; Indels 4; Gaps 3;

Qy 19 EOPTRYETLFOALDRNGDGVVDIGELGRNLGIPLODAEEKITFTGDVNVKDKLDFE 78

Db 73 EKERQIRDIRDLNDGIDIRDLTALKHETPHANLAPVIMSKSPDDEGRVDFY 132

Qy 79 EFMKYLKDKHEKMKLAFKSLDKNNKGIEASEIVOSLOTGLTISEQQAELILQSIDVG 138

Db 133 SFSSVYLENEOKLAEMFADNRHNDGLVDVEMKNYCKDIGVPLDDHKAQHVNMKMDQTG 192

Qy 139 TMTVDNEMRDYFLFNPVTDIIEIRFKHSTGIDIGSLTIPDEFTDEKKSGOMROL 198

Db 193 SASVDLEFEPFMMPLPSSDLKIDVDFWRHNLIIIDGDSQIPEDFSQEQHGIWRRHL 252

Qy 199 LAGIAGAVSTRTAPLDRUKIMQVHSGSKDKMNIFFGFRQVMEKGIIRSLWNGTGV 258

Db 253 VAGAGAVSTRTAPLDRUKIMQVHSGSKDKMNIFFGFRQVMEKGIIRSLWNGTGV 312

Qy 259 IKIAPETAVFWAYEYVKLLTTEE--GOKIGTFERFISGSMAGATAQFIYPMVEMKTRL 316

Db 313 IKIAPESAIKFCYDQLKRLIQKKNEBISTFERLCAGSAGASISQSIYPMVEMKTRL 372

Qy 317 AVGTGQYS-GIYDCAKKILKHGELGAFYGVNPLGIIPYAGIDLAVYELLKSYWLDN 375

Db 373 ALRKTGQDRGIHFHAKMTTREGICFCFYGLPNLGIIPYAGIDLAVYELLKSYWLDN 432

Qy 376 FAKDSVNPVGVNVLGGCALSSTCGQLASYPALVTRMQAAMLEGSPQLN-MVGLFERRI 434

Db 433 YETNSPSPGVALLACGCTSCSTCGSLSPFPFALVTRMQAAMLEGSPQLN-MVGLFERRI 492

Qy 435 ISKEGIPGLYRGITPDMFKVLPVAGISVYVYENMKOTILGV 474

Db 493 LQNSGVTFYRGITPDMFKVLPVAGISVYVYENMKOTILGV 532

RESULT 3

G89667

protein F17E5.2 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001

C:Accession: G89667

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; PMID:9906913; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G89667

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-531 <STO>

A:Cross-references: GB:chr\_X; PIDN:CAA90761.1; PID:q3876012; GSPDB:GN00028; CESP:F17E5

C:Genetics:

A:Gene: F17E5.2

A:Map position: X

Query Match 40.0% Score 991.5; DB 2; Length 531;

Best Local Similarity 42.3% Pred. No. 7.7e-58;

Matches 203; Conservative 92; Mismatches 138; Indels 47; Gaps 8;

Qy 16 QDAEOPTRYETLFOALDRNGDGVVDIGELGRNLGIPLODAEEKITFTGDVNVKDKLDFE 75

Db 11 QNISLSLRIRDMYDRLDADNDSIDIRLQAL-SLQAHIPASVAPKLLERMKSEHSORV 69

Qy 76 DFEPMKYLKDKHEKMKLAFKSLDKNNKGIEASEIVOSLOTGLTISEQQAELILQSID 135

Db 70 TYADFTNYVIAHEARLAEVFDKIDLSNGVDMAEIKSKYCKENGVLDDQKAMSVKMD 129

Qy 136 VDGTMVDNEMRDYFLFNPVTDIIEIRFKHSTG-----IDIGSL 178

Db 130 QSGSSSVNLNEFQDFMLLYPSTDMRDVDFWRHNLVCTLESNRPRTQNFOLIIDIGDG 189

Qy 179 TIPDEFTDEKKSGOMROLGAGIAGAVSTRTAPLDRUKI-----M 221

Db 190 QVPDFTQELLSGVWRRHLVAGVAGASRTCTAPDRIKYLVQVYLVHLLFHFHMLKA 249

Qy 222 MOVHSGSKDKMNIFFGFRQVMEKGIIRSLWNGTGVYKIAPETAVFWAYEYVKLLTTEE 281

Db 250 LQVNSTKTKNGVSVCHLLHAEGGKSPFRGNGINVIKIAPEASAKMFCYDQIKRMQOE 309

Qy 282 --EGOKIGTFERFISGSMAGATAQFIYPMVEMKTRLAVGKTCQY-SGIYDCAKKILKHE 338

Db 310 YKGAELSTIERLAGSSAGASISQSIYPMVEMKTRLALARTQOLKGMFHFHAKMYTKE 369

Qy 339 GLGAFYGVNPLGIIPYAGIDLAVYELLKSYWLDNFAKDSVNPVGVNVLGGCALSSTC 398

Db 370 GKIKYGLNPLGIIPYAGIDLVYELSKMT-TKYITTEHPGVALLACGCTSCSTC 428

Qy 399 QOLASYPALVTRMQAAMLEGSPQLN-----MVGLFRRISKEGIPGLYRGITPDMFK 453

Db 429 QOLASYPALVTRMQAAMLEGSPQLN-----SPKNSSTOPDMVGVGFKHILQTEGFTGLYRGITPDMFK 485

RESULT 4

T21074

hypothetical protein F17E5.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T21074

R:McMurray, A.

submitted to the EMBL Data Library, August 1995

A:Reference number: 219368

A:Accession: T21074

A:Status: preliminary

A:Molecule type: DNA

A:Map position: X

A:Gene: F17E5.2

A:Map position: X

A:Introns: 21/1; 48/3; 74/2; 111/3; 183/2; 384/1; 432/3; 503/1; 541/3

Query Match 39.8% Score 988.5; DB 2; Length 587;

Best Local Similarity 42.3% Pred. No. 1.4e-57;

Matches 202; Conservative 91; Mismatches 137; Indels 47; Gaps 8;

Qy 19 EOPTRYETLFOALDRNGDGVVDIGELGRNLGIPLODAEEKITFTGDVNVKDKLDFE 78